

# 統合ファブリック：データセンターインフラストラクチャを 簡素化する新しいテクノロジー

## 技術概要

概要	2
はじめに	2
従来のデータセンタートポロジ	2
10 Gigabit Ethernetを使用した統合ファブリック	3
iSCSI	3
FCoE (Fibre Channel over Ethernet)	4
データセンターでのFCoEによるネットワーク統合	5
統合ファブリックへの移行	5
データセンターブリッジング	6
統合ネットワークにおける仮想化	8
サーバー仮想化	8
NIC仮想化	8
バーチャルコネクタにおけるFlex-10の使用	8
まとめ	9
詳細情報	9
ご意見の送り先	9

## 概要

統合ファブリックは、データセンターの管理を簡素化するソリューションです。プロセス間通信 (IPC) とストレージに独自のファブリックを使用する、より複雑でコストのかかる手法ではなく、すべての通信 (サーバー、ストレージ、ネットワーク、リモート管理) を単一のファブリックに統合することにより、簡素化を実現します。本書では、HP が提供する新しい統合ファブリックテクノロジーについて説明します。ファブリックテクノロジーには、10GbE (10 Gigabit Ethernet) と FCoE (Fibre Channel over Ethernet) があります。さらに、10GbE と FCoE で複数のネットワークファブリックを結合して、単一の統合インフラストラクチャを構成する方法も含まれます。

## はじめに

複数のネットワークファブリックの実装、管理、コストに関する複雑さは、HP や他のベンダーによる統合ファブリックソリューション (単一ファブリックまたはファブリック統合とも呼ばれる) の研究を促す要因でした。ファブリックとは、1 つまたは複数のスイッチドネットワークノードで構成されるネットワークテクノロジーを指します。したがって、「ファブリック」とは一般的に個別のネットワークタイプを表すために使用される用語で、これには通信、ストレージ、管理、および高速ネットワークなどが含まれます。統合ファブリックは、データセンター管理の簡素化およびインフラストラクチャコストの削減を目的としたソリューションです。このソリューションは、IPC とストレージに独自のファブリックを使用する、より複雑でコストがかかる手法とは異なり、個別のネットワークタイプをすべて単一のファブリックに統合します。統合ファブリックソリューションは一般的なインターコネクトであるため、InfiniBand (IB) または 10GbE のいずれかを利用できます。HP の目標は、統合ファブリックテクノロジーにより、パフォーマンスを維持しながらインフラストラクチャと管理のコストを大幅に削減することです。本書では、統合ファブリック用の一般的なインターコネクトとして 10GbE がどのように使用されるのかを説明します。

## 従来のデータセンターポロジ

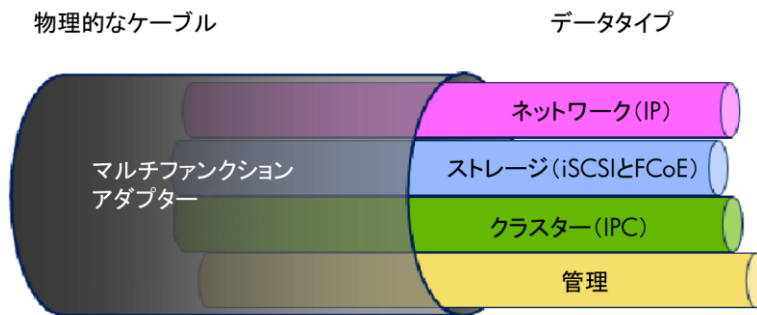
現在のデータセンターインフラストラクチャには、一般的に、容量が十分に活用されていない、リソースに柔軟性がなく単一用途向けである、管理にコストがかかる、といった特徴があります。その理由は、多くの場合、データセンターポロジには異なる種類のデータを管理するために別々の異機種ネットワークアプリケーションが含まれているからです。これらのアプリケーションそれぞれが、データセンターの複雑さ、コスト、管理の手間を増大させる要因となっています。複数のネットワークで、個別のスイッチ、各サーバー用のネットワークアダプター、および各ネットワークに対応するネットワーク管理システムが必要になります。

企業が、ストレージエリアネットワーク (SAN) のデータ、IPC クラスタリングのデータ、リモート管理のデータ、および Ethernet 通信のデータを管理するために 4 つの個別のネットワークをサポートするような場合もあるかもしれません。そのような場合、通常は、データの種類ごとに異なるネットワークとインターコネクトを使用するので、ネットワークアーキテクチャーの複雑さやネットワークファブリックを統合する際の複雑さが増大します。

## 10 Gigabit Ethernetを使用した統合ファブリック

統合ファブリックの基礎としてEthernetを使用する上での障壁の1つは、その帯域幅の狭さでした。10GbEテクノロジーが幅広く導入されるようになったことで、HPは、10GbEの帯域幅または低レイテンシの利点を必要とするアプリケーションのニーズに応えられる10GbEネットワークコンポーネントに期待しています。10GbEの登場により、すべてのデータセンターアプリケーションに対応する統合Ethernetスイッチングファブリックは、将来のデータセンターの統合およびアーキテクチャーの進化を支える基礎となることが期待されています。統合スイッチングファブリックとしてEthernetとインターネットプロトコル(IP)を使用することで、ネットワーク管理ツールを選択する際にもきわめて高い柔軟性が実現します。また、Ethernetの帯域幅が増加したことによって、より少ない数の物理リンクでより多くのデータを搬送できるようになっています(図1を参照)。

図1: すべてのタイプのトラフィックが同じリンクを共有



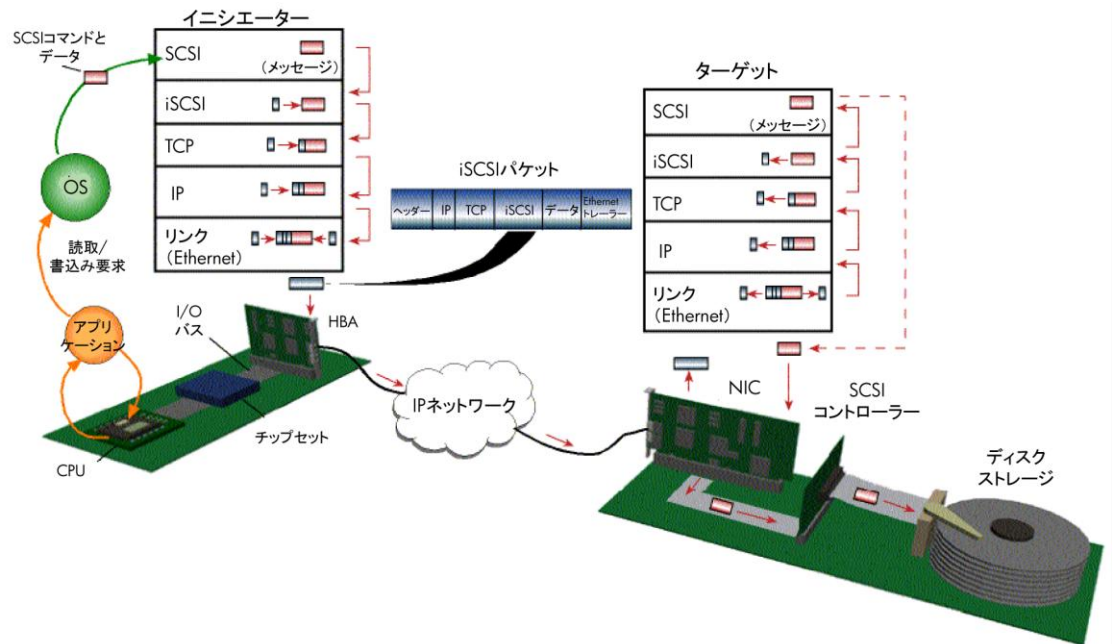
現在、統合ファブリック用として最も有望なのは、iSCSI(Small Computer System Interface)とFCoEの2つの伝送規格です。

### iSCSI

iSCSIは、イニシエーターとストレージリソース(またはターゲット)間のメッセージ交換に基づくSCSIアーキテクチャーモデルに従っています。iSCSIイニシエーターは、iSCSIプロトコルを使用してターゲットにアクセスします。通常、ターゲットはドライブエンクロージャーまたは別のコンピュータですが、テープドライブなど、iSCSIプロトコルをサポートする他のストレージデバイスになることもあります。パスの両端にあるiSCSIスタックでは、SCSIブロックコマンドをカプセル化してEthernetパケットを生成し、IPネットワーク経由で送信します。

図2は、イニシエーターとターゲット間のメッセージ交換を示しています。プロセスは、アプリケーションがオペレーティングシステム(OS)に対してデータの読み取りまたは書き込み要求を送信すると開始されます。OSは、メッセージの形で適切なSCSIコマンドとデータ要求を生成します。メッセージはIPネットワーク経由で送信される前にiSCSIによって処理され、要求がTCP/IP(Transmission Control Protocol/Internet Protocol)プロトコルスタックにカプセル化されます(ルーティング、エラーチェック、および制御情報を追加)。これは、ドライバーレベルまたはOSレベルのソフトウェアを使用して実行でき、ホストバスアダプター(HBA)にオフロードすることもできます。HBAはIPネットワーク経由でパケットを送信します。パケットがターゲットデバイスに到達すると、逆の処理を通してデータが再度組み立てられ、SCSIコントローラーに移動します。SCSIコントローラーはターゲットデバイスへのデータの書き込みまたはターゲットデバイスからのデータの読み取りを行うことで要求を実行します。読み取りトランザクションの場合、ターゲットはiSCSIプロトコルを使用してイニシエーターにデータを返送します。

図2: イニシエーターとターゲット間でのiSCSIプロトコルモードを使用したメッセージ交換



イニシエーターには、ソフトウェアイニシエーターとHBAがあります。ソフトウェアイニシエーターでは、プロトコルスタックを管理するためにCPUリソースが必要です。これよりも効果的な方法は、HPの内蔵NC373iマルチファンクションGigabitサーバーアダプターなどのiSCSI HBAに、プロトコルの管理をオフロードすることです。

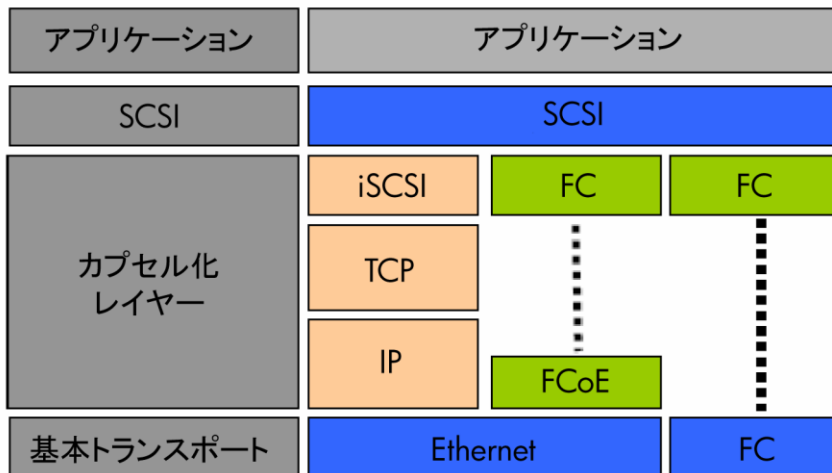
iSCSIを使用するとEthernetサーバーアダプターは、ネットワークアダプターおよびストレージHBAの両方として機能するようになります。このような柔軟性が備わっていることにより、1本の接続でストレージボックスとサーバーにアクセスでき、ネットワークインフラストラクチャが簡素化されます。HPのマルチファンクションアダプターは、iSCSI動作のハードウェアによる補助または高速化をサポートしています。つまり、iSCSIソフトウェアをサーバーアダプターにオフロードすることにより、システムパフォーマンスが向上します。

## FCoE(Fibre Channel over Ethernet)

FCoEは、情報技術規格国際委員会(INCITS)のT11技術委員会により策定された新しいテクノロジーです。FCoEはフロー制御に依存しており、バッファにほとんど空きがなくなったことを認識し、バッファに空きができて送信が再開できるようになるまで、送信側が送信停止を要求します。FCoEはEE(Enhanced Ethernet)に基づいているので、DCE(Data Center Ethernet)またはCEE(Converged Enhanced Ethernet)と呼ばれることもあります。IEEE(米国電気電子学会)は、ファイバーチャネルがEthernet上で効率的に動作するように、IEEE 802 Ethernet規格を強化する努力を続けています。Ethernetに対して最近実施された改善としては、ロスレスEthernetおよび一時停止方式のフロー制御の2つがあります。ロスレスEthernetは、パケットの消失を防止します。一時停止方式のフロー制御は、ネットワークが種類の異なるトラフィックを選択的に一時停止することを可能にします。

FCoEは、Ethernetファブリック内でファイバーチャネル(FC)フレームをカプセル化し、IPネットワークとして同じOSI(Open Systems Interconnection)レイヤーを使用します(図3を参照)。実際のアプリケーション環境では、ロジックにより、サーバーとネットワーク間で1つの共通FCoE接続を使用することが指示されます。FCoEは、現在利用されているのと同じ共通のドライバースタック、ケーブル接続、および管理アプリケーションを使用します。

図3: プロトコルスタックの比較



FCoEには、次のようないくつかの利点があります。

- FCoEは、FCドライバー、スイッチ、および他のインフラストラクチャを使用します。
- 既存のFCセキュリティおよび管理アプリケーションが変更されることはありません。
- FCoEは、データの消失が起こらない10GbEファブリックを使用します。
- 既存のSAN管理ツールを使用して、ストレージにアクセスできます。
- FCoEでは、トラフィックの優先度決定やフロー制御など、拡張されたEthernet機能を使用できます。

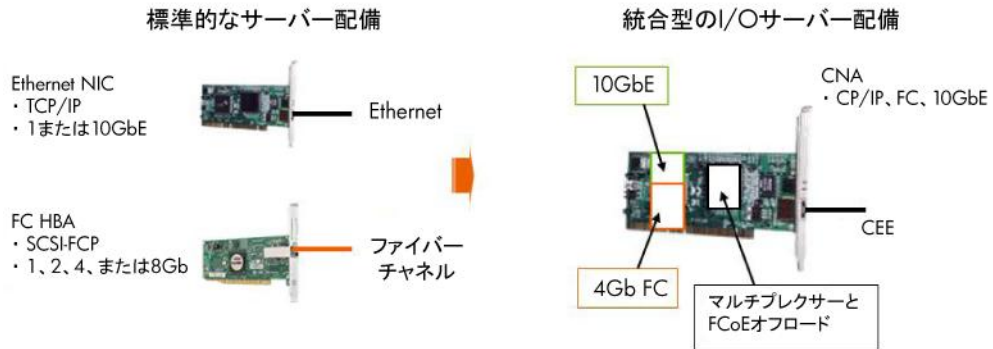
## データセンターでのFCoEによるネットワーク統合

HPは、FCoEによるネットワーク統合の実現の可能性に、積極的に取り組んでいます。FCoEには、同じ物理Ethernetレイヤーおよび現在利用されている共通の管理アプリケーションを使用することで、どのような規模のIT環境にも対応できる潜在的な可能性があります。

### 統合ファブリックへの移行

従来のデータセンターモデルでは、各サーバーで複数のHBAとNICを使用して、さまざまなネットワークで通信を行っています。統合モデルでは、サーバーに搭載されているHBAとNICを単一の統合ネットワークアダプター(CNA)に集約して、統合型のI/Oサーバー配備を実現します(図4を参照)。

図4: 標準的なI/Oサーバー配備と統合型のI/Oサーバー配備の比較



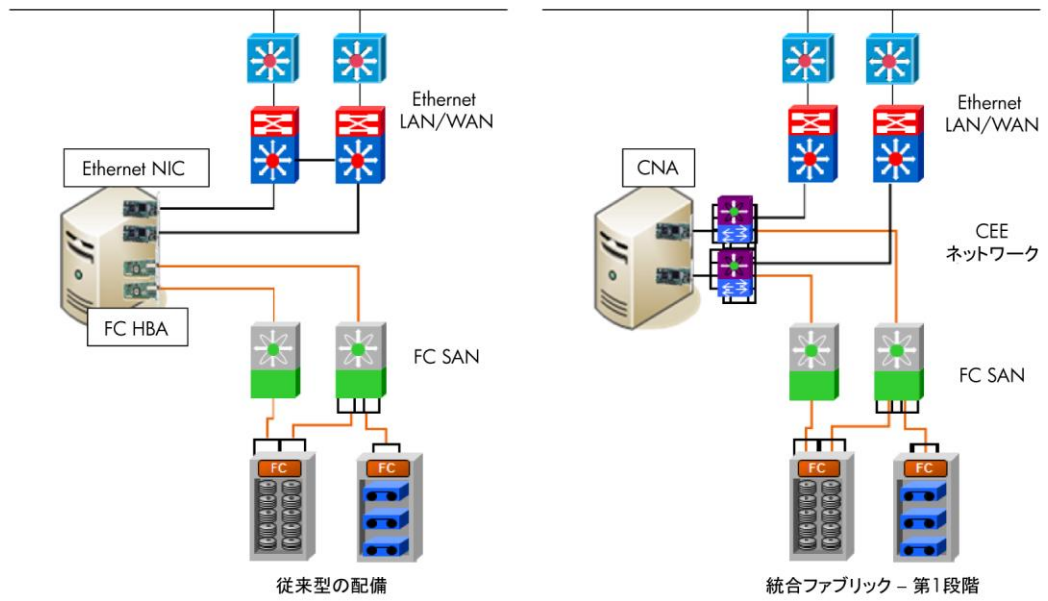
## データセンターブリッジング

DCEまたはCEEとも呼ばれるデータセンターブリッジング(DCB)は、Ethernetリンク内におけるトラフィックタイプおよび帯域幅共有の優先度を決定します。DCBは、統合ネットワークの配備で使用される既存のEthernet 802.1ブリッジ規格を次のように拡張して、すべてのアプリケーションを単一のEthernetリンクで動作させることを可能にします。

- 輻輳通知(CN): データの消失を減らすための輻輳管理を提供します。
- 優先度ベースフロー制御(PFC): 輻輳によるデータの消失を防止するリンクレベルのフロー制御メカニズムを提供します。
- 拡張伝送選択(ETS): 帯域幅をさまざまな種類のトラフィックに割り当てる共通の管理フレームワークを提供します。
- 検出および機能交換プロトコル: ネットワーク全体で一貫性のある設定を保証します。

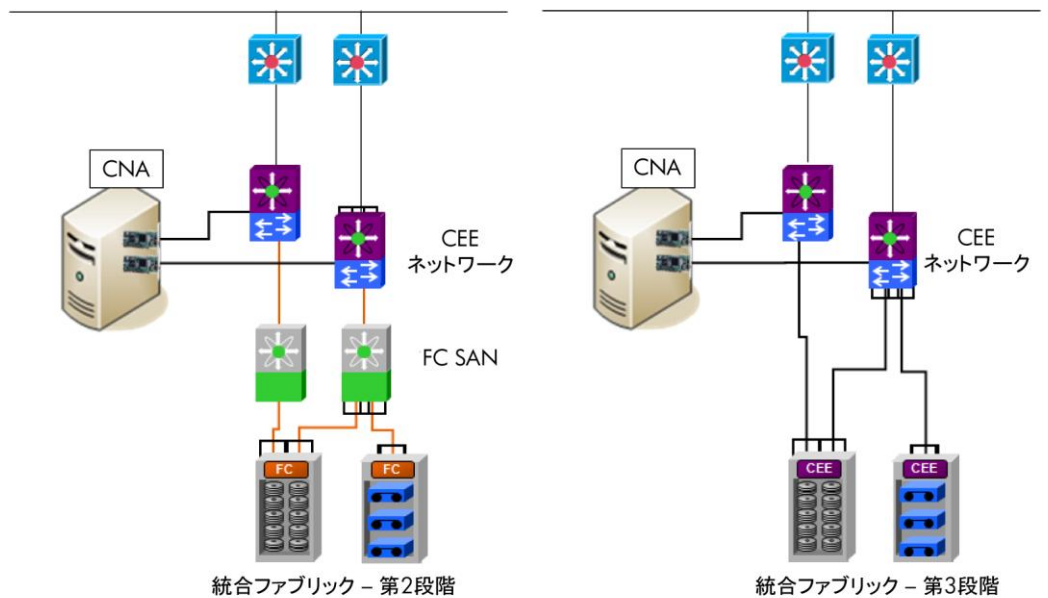
統合ファブリック、CNA、統合ファブリックブリッジへの移行の第1段階では、サーバー、LAN/WAN、およびSANの間で使用する接続が提供されます(図5を参照)。

図5: 従来型の配備と統合ファブリックの配備の第1段階



配備の次の段階では、既存のデータセンターを更新するか、または新しいデータセンターを構築します。最終的に、各サーバーで必要となるのは、リダンダントCNAの1つのペアのみになります。別々に設置されていたFC、10GbE、IBのスイッチは、統合ネットワークスイッチに置き換えられます(図6を参照)。

図6: 統合ファブリックの配備の第2段階と第3段階





## 統合ネットワークにおける仮想化

既存のコンピューティングリソース、サーバー、ネットワーク仮想化をより効果的に活用しようとする企業が急速に増加しています。HPバーチャルコネクト(VC)は、HP BladeSystem c-Classエンクロージャー用のインターコネクトモジュールとモジュールに組み込まれたソフトウェアで構成されます。このソフトウェアは、サーバー接続のセットアップと管理を簡素化します。HPバーチャルコネクトには、HP BladeSystem c-Class 1/10Gb バーチャルコネクト ファイバ イーサネットモジュール、HP BladeSystem c-Class 10/10Gb バーチャルコネクト Flex-10 イーサネットモジュール、HPバーチャルコネクトマネージャーが含まれています。HPバーチャルコネクトは、サーバーに搭載されているBladeSystem c-Class用のアダプターおよび新しい種類のEthernetインターコネクトモジュールを使用して、これらのサーバーNICのデータセンターファブリックへの接続を簡素化します。また、HPバーチャルコネクトは、Ethernet MACアドレスを安全に管理することにより、標準的なサーバーNICの機能を拡張します。

### サーバー仮想化

物理サーバー上にロードする仮想マシンの数が増えるほど、より多くの帯域幅を追加する必要があります。VMwareのベストプラクティスでは、仮想マシンを実行する物理サーバーのそれぞれに、6枚の1Gb NICを搭載することが要求されます。VMwareのこのガイドラインを使用した場合、仮想マシンをロードする物理マシンが2台あれば、1つの10Gb NICの性能を十分に活かします。仮想マシンテクノロジーの詳細については、<http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01067846/c01067846.pdf>にある技術概要の『Server virtualization technologies for x86-based HP BladeSystem and HP ProLiant servers』(英語)を参照してください。

### NIC仮想化

HP BladeSystem c-Classのサーバーブレードの場合、HP VC Ethernetモジュールを使用することで、IT管理者は、複数のモジュールを相互接続して、データセンターEthernetスイッチへのアップリンクを定義できるようになります。HP VC Ethernetモジュールでは、IT管理者は、各外部ネットワークにサーバーのどのNICポートを接続するかを選択できます。

### バーチャルコネクトにおけるFlex-10の使用

10GbE接続の帯域幅を十分に活用できるように、HPIは、HP BladeSystem c-ClassのアーキテクチャーにFlex-10テクノロジーを導入しました。Flex-10を使用することで、管理者は、1つの10Gbパイプラインの帯域幅を複数の「FlexNIC」に分割できます。分割された帯域幅のそれぞれは、10Gb接続の全体の帯域幅に対してユーザー定義部分として設定することができます。速度は、1秒当たり100メガビット～10ギガビットの範囲において、100メガビット単位で設定できます。10GbEパイプラインの分割には、次のような利点があります。

- 各サーバーのNIC接続の数を増やすことができます。これは、特に仮想マシン環境で重要です。
- ネットワーク機能に適した帯域幅を割り当てることができます(例: 仮想マシンの移行、管理コンソール、実務データ)。

Flex-10テクノロジーの詳細については、<http://h50146.www5.hp.com/products/servers/bladesystem/whitepaper/c01608922-J2/pdfs/c01608922-J2.pdf>にある技術概要の『HP Flex-10テクノロジー』を参照してください。



## まとめ

統合ファブリックは、サーバー、ストレージ、ネットワークトラフィックを共通のインターコネクトで統合することによってデータセンターの管理を簡素化する新しいテクノロジーです。HPは、10GbE接続を最大限に利用してデータセンターにおける統合ファブリックへの移行を促進するために、Flex-10などのインテリジェントでダイナミックなテクノロジーを導入しています。HPの目標は、データセンターに統合ファブリックを実装して、パフォーマンスを維持しながらインフラストラクチャと管理のコストを大幅に削減することです。

## 詳細情報

詳細情報については、下記の資料を参照してください。

資料	Webアドレス
HPマルチファンクションネットワーク製品	<a href="http://h18004.www1.hp.com/products/servers/proliant-advantage/networking.html">http://h18004.www1.hp.com/products/servers/proliant-advantage/networking.html</a> (英語)
HP ProLiantネットワーキング用のEthernetネットワークアダプター	<a href="http://h50146.www5.hp.com/products/servers/proliant/options/nic.html">http://h50146.www5.hp.com/products/servers/proliant/options/nic.html</a>
技術概要『Server virtualization technologies for x86-based HP BladeSystem and HP ProLiant servers』	<a href="http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01067846/c01067846.pdf">http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01067846/c01067846.pdf</a> (英語)
技術概要『10 Gigabit Ethernet technology for industry-standard servers』	<a href="http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01608915/c01608915.pdf">http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01608915/c01608915.pdf</a> (英語)
技術概要『HP Flex-10テクノロジー』	<a href="http://h50146.www5.hp.com/products/servers/bladesystem/whitepaper/c01608922-j2/pdfs/c01608922-j2.pdf">http://h50146.www5.hp.com/products/servers/bladesystem/whitepaper/c01608922-j2/pdfs/c01608922-j2.pdf</a>
HPバーチャルコネクトテクノロジー	<a href="http://h50146.www5.hp.com/products/servers/bladesystem/c/component/vc/">http://h50146.www5.hp.com/products/servers/bladesystem/c/component/vc/</a>

## ご意見の送り先

本書についてご意見がありましたら、[TechCom@HP.com](mailto:TechCom@HP.com)(英語)までご連絡ください。

© 2009 Hewlett-Packard Development Company, L.P.  
本書の内容は、将来予告なしに変更されることがあります。HP製品およびサービスに対する保証については、当該製品およびサービスの保証規定書に記載されています。本書のいかなる内容も、新たな保証を追加するものではありません。本書の内容につきましては万全を期しておりますが、本書中の技術的あるいは校正上の誤り、脱落に対して、責任を負いかねますのでご了承ください。

